

3/PRTS

09/857481

GR 98 P 5866
Version for foreign countries

JC18 Rec'd PCT/PTO 04 JUN 2001

Description

Method and arrangement for reducing a number of measured values of a technical system

5

The method relates to a method and an arrangement for reducing a number of measured values of a technical system.

10 A technical system, for example a process engineering plant, supplies a large quantity of measured values per unit time, using different measured-value pickups (sensors). In the course of a number of days or weeks, a quantity of data accumulates which requires a
15 correspondingly high computational power to process it. If the measured values are to be used to adapt or reset the technical system, operations are often necessary whose complexity admits only some of the measured values. It is then a great disadvantage to select a
20 specific number of measured values arbitrarily from the quantity of all measured values and to process them further, since measured values with low significance, for example measured values with a high degree of interference, have a considerable influence on the
25 overall result and falsify the latter considerably.

The object of the invention is to specify a method and an arrangement for reducing measured values, it being ensured that the remaining measured values have a high
30 significance with regard to their description of the technical system.

This object is achieved in accordance with the features of the independent patent claims. Developments of the
35 invention emerge from the dependent claims.

FOR "59" 5360

In order to achieve the object, a method of reducing a number of measured values of a technical system is specified, in

09857481.060401

which the measured values are divided into classes in accordance with predefined criteria. The measured values in a class are assessed and those measured values whose assessment lies below a predefined first
5 threshold value are screened out.

Screening out the measured values results in a reduction with regard to the number of measured values. A considerably reduced number of measured values is
10 thus available for further processing. The further processing can be carried out with a lower computational outlay as compared with the non-reduced number of measured values.

15 One development consists in that the classes themselves are also assessed. In particular, a class whose assessment lies below a predefined second threshold value can be screened out. This results in an additional reduction in the number of measured values.

20 Another development consists in that one criterion for the division into classes consists in that, for each class, measured values relating to a predefinition of setting parameters of the technical system are
25 determined. Typically, the technical system is set using a predefined number of setting parameters and after the setting has been carried out, there follows a (mostly time-delayed) reaction of the system to the setting parameters (transient response, transient
30 process of the system). After the setting operation, a specific quantity of measured values which can be associated with the transient process are therefore picked up, measured values which are associated with the predefined set of setting parameters continuing to
35 accumulate after the transient process has concluded (transition to steady-state operation). By adjusting the setting parameters, a new class is determined. All the measured values which respectively accumulate after an

RECEIVED FEBRUARY 1998

adjustment to the setting parameters belong in their own class.

09857481 050401

Another development consists in that measured values within a class which can be associated with the respective transient process are screened out. In addition, erroneous measured values can be screened out. In many cases, the setting of large technical systems depends on long-term, steady-state operation. Measured values which relate to the transient process (of short duration in comparison with the steady-state operation after the transient process has concluded) are expediently screened out, since measured values for the steady-state operation are falsified by them. In particular within the context of modeling the technical system, the measured data about the steady-state behavior of the technical system are of interest.

15

One refinement consists in reducing the number of measured values in a class in that at least one representative value for the measured values in the class is determined. Such a representative value may be:

20

- a) an average (e.g. a sliding average) of the measured values of the class,
- b) a maximum value of the measured values of the class,
- 25 c) a minimum value of the measured values of the class,
- d) a median.

In the case of variant d), one advantage resides in the fact that it is always possible to determine a value which is actually present, while the average a) does not itself occur as a value.

30

Depending on the application, a suitable selection for determining the representative value of a class may be made.

35

"060404" "060404" "060404"

An entire class with measured values can be screened out if said class contains less than a predefined number of measured values.

- 5 Another refinement consists in that those measured values are screened out which differ from a predefinable value by more than a predefined threshold value.

09857491 060401

The predefinable value can be an average of all the measured values of the class or a measured value to be expected in response to the respective setting parameters of the technical system.

5

A development consists in that the measured values which are reduced in number are used for the simulation and/or the draft design of this or another technical system.

10

The draft design of the technical system can in this case comprise both adaptation and redrafting of this or another technical system.

15

In addition, within the context of the simulation, the behavior of the technical system can be checked by using the reduced number of measured values, with the objective of altering the system for developing a new system with altered predefinitions.

20

In addition, in order to achieve the object, an arrangement for reducing a number of measured values of a technical system is specified which has a processor unit, which processor unit is set up in such a way that the measured values can be divided into classes in accordance with predefined criteria. Measured values in a class can be assessed, measured values whose assessment lies below a predefined first threshold value are screened out.

30

This arrangement is particularly suitable for carrying out the method according to the invention or one of its developments explained above.

35

Exemplary embodiments of the invention will be illustrated and explained below using the drawing, in which

"PATENT" 25866

Fig. 1 shows a block diagram which contains steps in a method for reducing measured values;

5 Fig. 2 shows a schematic sketch of a recovery boiler;

Figs. 3-5 show input variables, actuating variables and outward variables of the recovery boiler.

10 Fig. 1 shows a block diagram which contains steps in a method of reducing measured values. In a first step 101, the measured values are divided into classes. In particular, the division into classes is carried out with the effect that all the measured values which
15 belong to one set of setting parameters are combined in one class. To this extent, each alteration to the setting parameters of the technical system found a new class. In particular, altering the setting parameters is associated with a transient process of the technical
20 system, this process, as opposed to a steady-state behavior, containing extreme fluctuations in the measured values. In a step 102, individual measured values within one class are screened out. These may be, for example, erroneous measured values, that is to say
25 measured values which exhibit a high deviation with respect to the other measured values or an average of the measured values or measured values from the transient process. There are a number of possible ways of screening out individual measured values within a
30 class:

1. Measured values which are too poor (based on a predefined comparative value);
2. Measured values which belong to a transient process;
- 35 3. Measured values which are erroneous;
4. Determining a representative measured value as a representative for a plurality of measured values, in that the representative measured value is determined as an average of the

09357441 060401

GR 98 P 5866

- 5a -

measured values in a class or as a maximum value or a minimum value of these measured values.

09857431.030431
"FORM 90" T 2475860

Measured values of this type are preferably not taken into account; they are removed from the respective class. This results in a considerable reduction in the number of measured values. In a step 103, individual classes are screened out. One criterion for screening out an entire class consists in that the class contains less than a predefined number of measured values. In a step 104, the measured values reduced in number are used for further processing. Further processing is, in particular, a simulation and/or a draft design of the technical system.

Fig. 2 shows a schematic sketch of a recovery boiler. In the following text, by using the example of a "recovery boiler", an exemplary embodiment of the method described above will be illustrated.

In the paper and pulp industry, various chemicals and also heat and electrical power are needed for the digestion of pulp. With the aid of the recovery boiler, the chemicals used and additional thermal energy may be recovered from a thickened waste process liquor (black liquor). A measure of the recovery of the chemicals is of critical importance for the economy of the overall plant.

The black liquor is burned in a char bed 201. In the process, an alkaline melt is produced and flows away via a line 202. In further process steps, the chemicals used are recovered from the constituents of the alkaline melt. Heat of combustion which is released is used to generate steam. The combustion of the waste liquor and therefore the recovery of the chemicals begins with the atomization of the black liquor via atomizer nozzles 204 into a combustion chamber 203. As they fall through the hot flue gas, particles of the atomized black liquor are dried. The dried liquor particles fall onto the char bed 201, first combustion and

40

09857481 060401

chemical reduction taking place. Volatile constituents and reaction products pass into an oxidation zone, in which oxidizing reactions proceed and in which the combustion is completed.

5

Important objectives for the control of the recovery boiler are the steam production in order to obtain power, compliance with emission values from environmental points of view and the efficiency of the chemical reduction.

10

The combustion operation, and therefore the objectives, are controlled in particular by the supply of air at three levels (primary air (PA), secondary air (SA), tertiary air (TA)). The overall process is subject to numerous influences, which have to be taken into account during the modeling:

15

- a) the measurement of the variables is subject to fluctuations which are often extreme;
- b) influencing variables which are not measured and cannot be measured exist;
- c) at each alteration to the settable parameters transient processes occur;
- d) the technical plant becomes soiled and is cleaned at predefined intervals, which has the effect of a drift over time in each case in the system behavior.

20

25

The measured variables of the overall process are subdivided into input variables (cf. Fig. 3) and output variables (cf. Fig. 5). Measured variables are stored every minute. Four of the input variables are simultaneously also actuating variables (also: settable parameters; cf. Fig. 4). The actuating variables are to be viewed substantially as free parameters of the overall process which can be set independently of one another. Some of the other

30

35

09857481 060404

input variables are more or less dependent on the actuating variables. According to one predefinition, the variables "BL Front Pressure" and "BL Back Pressure" are always to be regulated equally in the recovery boiler. The four actuating variables (cf. Fig. 4) are preferably to be stored as actuating variables (with the desired, preset value) and as input variables (with the measured, actual value).

10 In the recovery boiler, one problem consists in the fact that, depending on the settable parameters, specific objectives, which are defined via measured variables, have to be met. Here, a three-stage procedure is selected in order to solve the problem:

15

1. The objectives to be considered are modeled by means of stochastic methods, these models being updated by means of new measurements (data-driven, empirical modeling). In this case, it is expedient to use not just a single model but global models for the identification of interesting areas in a parameter space determined by the objectives, and to use local models for the exact calculation of optimum operating points. The models used are assessed by means of quality measures.

20

25

2. If the models considered are not sufficiently accurate because of the state of the data (quality measure), new operating points are deliberately evaluated in order to improve the model (experimental design). In addition, by using global stochastic optimization methods with regard to the objectives, attractive regions are identified on the basis of the current global model.

30

35

3. For the local optimization, local models are constructed, and the data sets which are available

09857481 060401
T04090 "T842560

are, if appropriate, deliberately expanded
(experimental design).

0957431 050401
T04050 T845860

The objectives are physical/technical or economic criteria which, as a rule, have to meet boundary conditions and/or safety conditions. It is often the case that a number of these criteria have to be considered at the same time. A stochastic model can be used in particular for the purpose of simulating the objectives to be optimized and their dependence on the parameters to be set in the computer. This is necessary when measurements are very costly or very time-consuming. In the case of safety requirements, possible hazardous situations can be avoided.

In the case of the recovery boiler, on-line optimization, which is based on a plurality of items of data, is necessary, since the physical/chemical processes cannot be modeled quantitatively with sufficient accuracy and because the behavior of the plant is subject to fluctuations in the course of operation. The knowledge about this behavior must continually be expanded by means of the deliberate selection of new operating points. Therefore, within the context of on-line optimization, the above-described three-stage procedure of stochastic modeling at mathematical optimization is to be recommended.

DESCRIPTION OF THE INPUT VARIABLES

The a input variables ($a \in \mathbf{N}$, \mathbf{N} : set of natural numbers) generally depend on n actuating variables $n \in \mathbf{N}$ and on random effects. They can be described as follows:

Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space and \mathcal{B}^v be a Borel σ -algebra over \mathbf{R}^v (\mathbf{R} : set of real numbers) for each $v \in \mathbf{N}$. The input variables are represented by a projection φ which can be measured via $\mathcal{B}^n \times \mathcal{S} - \mathcal{B}^a$:

$$\varphi: \mathbf{R}^n \times \Omega \rightarrow \mathbf{R}^a \quad (1).$$

The definition set of the projection φ is a cartesian product of two sets. If one considers the respective projections onto the individual sets, then the following projections are obtained:

5

$$\varphi_x: \Omega \rightarrow \mathbb{R}^a, \omega \rightarrow \varphi(x, \omega) \quad \text{for all } x \in \mathbb{R}^n \quad (2),$$

$$\varphi^\omega: \mathbb{R}^n \rightarrow \mathbb{R}^a, x \rightarrow \varphi(x, \omega) \quad \text{for all } \omega \in \Omega \quad (3).$$

$\{\varphi_x; x \in \mathbb{R}^n\}$ is a stochastic process having an index set \mathbb{R}^n and a projection φ^ω is a path in this stochastic process for each event $\omega \in \Omega$.

10

In the case of the recovery boiler, $n=4$ and $a=14$ (following the elimination of the variable "BL Back Pressure").

15

Because of the required ability to measure the projection φ_x for each $x \in \mathbb{R}^n$, the projection φ_x is a random variable. Under suitable additional preconditions, expected values and higher moments can be considered. This access makes the step possible from stochastic models to deterministic optimization problems. In the case of a deterministic optimization problem, the target function can be set directly by means of a variable, while the stochastic variable influences the target function but does not permit any deliberate setting.

25

DESCRIPTION OF THE OUTPUT VARIABLES

30

The process model M of the recovery boiler will be described as a function depending on the input variables and further random effects. In this case, let $(\Omega, \mathcal{S}, \mathcal{P})$ be the above probability space. The process model M is then a projection which can be measured by $\mathbb{B}^a \times \mathcal{S} - \mathbb{B}^b$:

35

$$M : \mathbf{R}^a \times \Omega \rightarrow \mathbf{R}^b \quad (4),$$

where b refers to the number of output variables.

5 Since the recovery boiler is subject to a cyclic drift over time (from cleaning phase to cleaning phase), a description using a time parameter is also conceivable. The output variables may be represented by projections ψ that can be measured by $\mathbf{B}^n \times \mathbf{S} - \mathbf{B}^b$:

$$\psi : \mathbf{R}^n \times \Omega \rightarrow \mathbf{R}^b \quad (5),$$

$$10 \quad (x, \omega) \rightarrow M(\psi(x, \omega), \omega) \quad (6).$$

If the respective projections onto the individual sets of the definition set are considered, then the following projections are obtained

15

$$\psi_x : \Omega \rightarrow \mathbf{R}^b, \omega \rightarrow \psi(x, \omega) \quad \text{for all } x \in \mathbf{R}^n \quad (7),$$

$$\psi^\omega : \mathbf{R}^n \rightarrow \mathbf{R}^b, x \rightarrow \psi(x, \omega) \quad \text{for all } \omega \in \Omega \quad (8).$$

$\{\psi_x; x \in \mathbf{R}^n\}$ is a stochastic process having an index set \mathbf{R}^n , and the projection ψ^ω is a path in this stochastic process for each $\omega \in \Omega$.

20

In the recovery boiler, $b=15$.

The fact that, when defining ψ , no distinction is drawn between the events ω used, does not mean that there is any restriction, since Ω can be represented as a cartesian product of an Ω_1 and an Ω_2 . The above representation therefore also comprises the model:

25

$$\psi : \mathbf{R}^n \times \Omega_1 \times \Omega_2 \rightarrow \mathbf{R}^b \quad (9),$$

$$(x, \omega_1, \omega_2) \rightarrow M(\varphi(x, \omega_1), \omega_2) \quad (10).$$

DESCRIPTION OF THE DATA SETS AVAILABLE

5 Using the descriptions in the two preceding sections, it is possible to combine the input variables and the output variables together to form measured variables Φ . Φ is a projection that can be measured by $\mathbf{B}^n \times \mathbf{S} - \mathbf{B}^m$, where $m = a + b$, and

10

$$\Phi : \mathbf{R}^n \times \Omega \rightarrow \mathbf{R}^m \quad (11),$$

$$(x, \omega) \rightarrow \begin{pmatrix} \varphi(x, \omega) \\ \psi(x, \omega) \end{pmatrix} \quad (12).$$

If the respective projections onto the individual sets of the definition set are considered again, then the
15 following projections are obtained:

$$\Phi_x : \Omega \rightarrow \mathbf{R}^m, \omega \rightarrow \Phi(x, \omega) \quad \text{for all } x \in \mathbf{R}^n \quad (13),$$

$$\Phi^\omega : \mathbf{R}^n \rightarrow \mathbf{R}^m, x \rightarrow \Phi(x, \omega) \quad \text{for all } \omega \in \Omega \quad (14).$$

$\{\Phi_x; x \in \mathbf{R}^n\}$ is a stochastic process with an index set \mathbf{R}^n and the projection Φ^ω is a path in this stochastic
20 process for each $\omega \in \Omega$.

For each chosen tuple x of actuating variables, a large number of implementations of Φ_x in the recovery boiler are determined and stored, that is to say for each $x_j \in$
25 \mathbf{R}^n , numerous implementations

098544 "0904" 04425860

$$\Phi_{jk} := \Phi(x_j, \omega_{jk}) \quad (15)$$

with $\omega_{jk} \in \Omega$; $k = 1, 2, \dots, v_j$;

$v_j \in \mathbf{N}$; $j = 1, 2, \dots, u$; $u \in \mathbf{N}$

are considered. The stored data sets D_{jk} of the recovery
5 boiler are therefore $(n + m)$ tuples:

$$D_{jk} = \begin{pmatrix} x_j \\ \Phi_{jk} \end{pmatrix}, \quad k = 1, 2, \dots, v_j; \quad j = 1, 2, \dots, u \quad (16).$$

Here, $D_{j_1 k_1}$ is stored before $D_{j_2 k_2}$ if

10

$$(j_1 < j_2) \vee ((j_1 = j_2) \wedge (k_1 < k_2))$$

DATA COMPRESSION BY DIVIDING THE CLASSES OF PARAMETERS

15 Since, for each tuple x of actuating variables, there
are generally a number of implementations of Φ_x ,
because of the complex stochastic properties of the
process to be considered, the first step in the
statistical data analysis is obviously to divide the
20 classes of parameters by forming arithmetic averages.
In addition, obviously erroneous data sets are
separated out. An obviously erroneous data set is, for
example, a physically impossible measurement which
cannot possibly occur in real terms, in particular on
25 the basis of a setting which has been made.

Procedure:

- 30 1. Data sets in which the variable "BL Front
Pressure" is not equal to the variable "BL Back
Pressure" are

screened out, since these two values must be equal according to the predefinition of the plant control system. The loss of data is very small.

- 5 2. The data sets are divided into classes in which the four setting parameters (PA, SA, TA, BL Front Pressure, see above) are successively constant over time, that is to say the j th class consists of the data sets D_j .
- 10 3. Classes in which there are fewer than 30 data sets are screened out, in order that transient processes do not have any great influence.
- 15 4. For each class, an arithmetic average $\bar{\Phi}_j$ and an empirical standard deviation s_j are determined for all the measured variables:

$$\bar{\Phi}_j = \frac{1}{v_j} \cdot \sum_{k=1}^{v_j} \Phi_{jk} \quad (17),$$

$$s_j = \begin{pmatrix} \left(\frac{1}{v_j - 1} \cdot \sum_{k=1}^{v_j} (\Phi_{jk}^{(1)} - \bar{\Phi}_j^{(1)})^2 \right)^{\frac{1}{2}} \\ \vdots \\ \left(\frac{1}{v_j - 1} \cdot \sum_{k=1}^{v_j} (\Phi_{jk}^{(m)} - \bar{\Phi}_j^{(m)})^2 \right)^{\frac{1}{2}} \end{pmatrix} \quad (18).$$

- 20 5. Classes in which the averages for the variables PA, SA, TA or BL Front Pressure are too far removed from the corresponding setting parameters are screened out. In these classes, therefore, the
- 25 setting values could not be reached.

09857441-050401

CHARACTERISTIC STATISTICAL VARIABLES FOR THE GIVEN CLASSES AND THEIR GRAPHICAL REPRESENTATION

In addition to the arithmetic averages and the empirical standard deviations which have been determined for the individual classes, a common standard deviation s is further determined in accordance with

$$s = \left(\begin{array}{c} \left(\frac{1}{v-1} \cdot \sum_{j=1}^u (v_j - 1) s_j^{(1)2} \right)^{\frac{1}{2}} \\ \vdots \\ \left(\frac{1}{v-1} \cdot \sum_{j=1}^u (v_j - 1) s_j^{(m)2} \right)^{\frac{1}{2}} \end{array} \right) \quad (19)$$

10

Here, u stands for the number of classes (205 here) and v for the sum of v_j , that is to say v is the number of all the measured values used (38,915 here).